SDN 场景中基于双向流量特征的 DDoS 攻击检测方法 *

陈 超，曹晓梅

(南京邮电大学 计算机与软件学院，南京 210000)

摘 要：传统网络资源的分布式特性使得管理员较难实现网络的集中管控，在分布式拒绝服务攻击发生时难以快速准确地检出攻击并溯源。针对这一问题，结合软件定义网络集中管控、动态管理的优势和分布式拒绝服务攻击特点，首先引入双向流量概念，提出了攻击检测四元组特征，并利用增长型分层自组织映射算法对网络流中提取的四元组特征向量快速准确地分析并分类，同时提出了一种通过自适应改变监控流表粒度以定位潜在受害者的检测方法。仿真实验结果表明，提出的四元组特征及下发适量监控流表项的检测算法能以近似 96%的准确率检出攻击并定位受害者，且对控制器造成的计算开销较小。

关键词：软件定义网络；双向流量；四元组特征；分布式拒绝服务攻击；增长型分层自组织映射

中图分类号：TP393 doi: 10.3969/j.issn.1001-3695.2018.01.0039

Distributed denial of service attack detection based on bidirectional traffic feature in software defined network

Chen Chao, Cao Xiaomei

(*School of Computer & Software*, *Nanjing University of Posts & Telecommunications*, *Nanjing* 210000, *China*)

Abstract: The distributed nature of traditional network resources makes it more difficult for administrators to realize the centralized control of the network. It is difficult to quickly and accurately detect and trace the DDoS attacks when distributed denial of service attacks occur. To solve this problem, combined with the advantages of centralized management and control of software defined network, the advantages of dynamic management and the characteristics of DDoS attacks, this paper first introduced the concept of bidirectional traffic feature, put forward the four-tuple characteristics of attack detection and made use of the growth hierarchical self-organizing map algorithm to analyze and classify the quaternion eigenvectors extracted from network flows quickly and accurately. At the same time, this paper proposed a new detection method that locates potential victims by adaptively changing the granularity of flow table. Simulation results show that the four-tuple features, as well as the detection algorithm issuing the monitoring flow entry, can detect DDoS attacks and pinpoint the victim with accuracy of nearly 96%, and the computational overhead for the controller is small.

Key words: software defined network; bidirectional traffic feature; four-tuple; distributed denial of service; growing hierarchical self-organization map

0 引言

软件定义网络(software-defined networks,SDN)技术可以使网络管理员根据网络实际状况动态更新网络流规则，极大地简化了网络管理的复杂性[1]。得益于这种集中式的网络管控能力，学术界和工业界对于 SDN 的研究逐年升温。然而，SDN 网络其特有的转控分离的架构模式也引入了更多的安全威胁。据卡巴斯基发布的 2017 安全报告显示，DDoS 攻击较去年在攻击次数和持续时间上均呈上升态势[2]，而传统网络与 SDN 网络的融合是未来发展的趋势，因此 SDN 场景下的 DDoS 攻击检测成为近些年研究的热点。

文献[3]将 SDN 场景中的 DDoS 攻击分为 OpenFlow 交换机流表攻击和 SDN 控制层带宽攻击两类。本文主要研究后一种攻击，而利用 SDN 控制器集中管控和可编程特性对网络流量进行审查分析是其经典的检测方法。

Braga 等人[4]利用 SDN 控制器的可编程特性和 SOM 算法分类计算时计算量小的优势，提出了一种基于流量特征属性的 DDoS 攻击检测方案。实验结果表明，利用流量特征属性检测

收稿日期：2018-01-22；**修回日期：**2018-03-19 **基金项目：**国家自然科学基金资助项目（61202353）；国家“973”计划资助项目（2011CB302903）；江苏高校优势学科建设工程资助项目（yx002001）

作者简介：陈超（1992-），男，江苏南京人，硕士，主要研究方向为信息安全（1584621666@qq.com）；曹晓梅（1974-），女，副教授，博士，主要研究方向为计算机网络、信息安全.

DDoS 攻击较传统检测方法开销更小，检出率更高。然而，由于实际网络场景中流量状况错综复杂，SOM 对其自身结构的初始权值的选取十分敏感，所以针对不同网络状况较难选取最优初始参数。同时，由于仿真实验网络规模的限制，文章默认用于统计 SDN 流量的流表项总是可以成功安装于 SDN 交换机中，并未考虑 SDN 交换机流表存储空间大小的限制，这在一定程度上限制了文章算法的适用性。

Mousavi 等人[5]通过统计恒定数量网络数据包目的 IP 地址的熵值，并判断其是否小于指定阈值的方法来确定当前 SDN 网络是否遭受 DDoS 攻击。虽然此检测方法具有开销低的优点，但在阈值的确定上需要提前进行大量的样本实验或依靠专家经验；与此同时，DDoS 攻击种类众多，在网络流量特征上呈现出多样性的特点，因此仅选取单一的特征属性显然不能保证 DDoS 攻击的高检出率。针对这一问题，Yan 等人[6]将流请求速率、目的 IP 地址熵值、目的端口熵值和源 IP 地址熵值这四种流量特征作为评价因素集，利用模糊数学中的模糊综合评判方法对不同程度的 DDoS 攻击进行模糊评价。然而模糊综合评判方法中评价因素的初始最优权重较难确定，且对于网络场景的变更难以做到自适应，因此该检测方法有进一步改进的余地。

姚琳元等人[7]提出了基于网络对象目的 IP 地址的检测七元组，并利用增长型分层自组织映射(growing hierarchical self-organization map,GHSOM)算法对收集到的七元组特征向量进行分析分类处理，以有效检出 DDoS 攻击受害者。实验结果表明，所提七元组的加权检测率接近 90%，且检测算法对控制器的开销基本可忽略。然而文章未考虑收集网络流量过程中对于 SDN 交换机流表空间大小的影响，且该方法较难区分高速率正常突发网络流量和 DDoS 攻击异常流量。

武泽慧等人[8]构建了一种 SDN 场景下基于 OpenFlow 的交换机洗牌模型，通过在交换机上部署流量吞吐率检测器，一旦发现网络包吞吐率大于指定阈值即判定网络正遭受 DDoS 攻击，进而在控制端触发交换机洗牌算法以分离攻击者和合法用户。仿真实验表明，与传统已有检测工具相比，本文方法在攻击流筛选能力上更胜一筹，且降低了误报率，但也一定程度上增加了数据包转发的延迟。与此同时，网络状况良莠不一，当发生高速率正常突发流时也会触发洗牌算法，进而消耗控制器系统资源。

通过上述描述分析可知，提出一种 SDN 网络下合理的 DDoS 攻击检测方案需要综合考虑多方面因素，如 DDoS 攻击检测算法的分类精度与实时性、算法对于控制器和交换机资源的消耗，以及对于高速率突发流等特殊网络状况的处理策略等。本文受文献[4]的启发，引入双向流特征的思想，在发生 DDoS 攻击时，对于受害者而言，网络流进与出之间的流速率相差较大，若不考虑双向速率差异，则会将网络中的高速率正常突发流量判定为异常流量，进而造成用户正常访问的阻断。然而，由于 SDN 交换机流表存储空间是有限的[9]，所以流表的有限性存储也应当作为 DDoS 检测重点考虑的因素。同时，GHSOM 算法近些年因其结构可拓展性强、分类精度高等优势在入侵检测中的应用范围逐步扩大，因此本文结合 SDN 高效下发流表及 GHSOM 算法分类优势，提出了一种 SDN 场景下基于双向流量特征的 DDoS 攻击检测方案。

本文特色在于：a).合考虑 SDN 架构及 DDoS 攻击特点，为准确检出 DDoS 攻击，本文引入双向流概念，提出了基于目的 IP 地址的 DDoS 检测四元组特征；b)考虑到 OpenFlow 交换机流表空间大小的限制，本文提出了一种流表项监控粒度自适应调整的策略；c)利用 GHSOM 算法对 SDN 网络中提取到的四元组特征向量分析并分类，之后在基于 Ope-nDayLight 的 SDN 仿真平台上分析了所提检测算法的可行性及检测性能。

1 分类算法 GHSOM 介绍

近些年来 SOM 算法应用广泛[10]，然而算法在实际运行过程中，存在结构不可动态改变、学习过程较为复杂等缺陷，因此近些年来专家学者对 SOM 的改进算法 GHSOM[11]进行了较多研究，其步骤大致可分为三个阶段：

a)初始化阶段。第 0 层包含单个神经元节点，其权值矢量 m_0 被初始化为所有输入向量的期望值，进而计算层 0 的平均量化误差（mqe_0）：

$$mqe_0 = (1/n)*\|m_0 - x\| \tag{1}$$

其中：x 表示输入向量；n 表示输入向量的个数。

b)网络结构训练和成长阶段。在竞争学习原则下，当且仅当 $mqe_0 \geq \Gamma_m mqe_0$ 时，神经元横向扩展。同时找到具有最大量化误差的神经元节点，并在此节点、与之相差最大的邻居神经元节点之间插入一行或一列并计算其权向量，若不满足上述条件，则进行纵向拓展。其中 Γ_m 为横向拓展系数。

c)展开或终止分层阶段。神经网络训练完成后，考察神经元是否符合全局终止条件，即 $mqe_i < T_u \cdot mqe_0$，对于不符合条件的神经元，重新进行步骤 b)，直至各层神经元都满足全局终止条件为止。

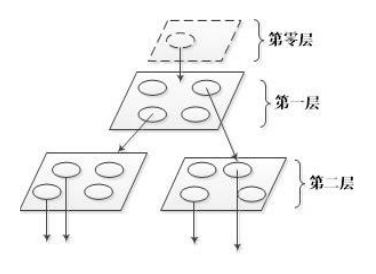


图 1 GHSOM 扩展图

2 DDoS 分类检测四元组特征

为了精确检出 DDoS 攻击，系统需要提取简单高效的网络流量特征。在 DDoS 攻击发生时，对于受害者而言，其接收的网络数据包无论在大小上还是数量上均异常增长；与此同时，在 DDoS 攻击发生时受害者端进出双向之间的数据包特征差异较明显，因此本文引入了包括网络流和双向流两类流量特征，

如表 1 所示，网络流特征包括 D_{ip}_PC 和 D_{ip}_BC；双向流特征包括 D_{ip}_PCA 和 D_{ip}_BCA。

本文将周期性提取监控流表项所审查 IP 网段的四元组网络特征信息，而一旦系统检出 DDoS 攻击受害者，则可判定网络正在遭受 DDoS 攻击。假定现行 SDN 网络中存在网络流 f_{ji}，代表从源地址段 j 到目的地址段 i 的网络流，定义目的地址段范围 i 的 IP 地址数目为 $Gra(i)$，则提取的 DDoS 攻击受害者网络流四元组信息用公式表示如下：

$$D_{ip}_BC = \sum Byte_{ji} / Gra(i) \tag{2}$$

$$D_{ip}_PC = \sum Pk_{ji} / Gra(i) \tag{3}$$

$$D_{ip}_PCA = (\sum_j Pkt_{ji} - \sum_j Pkt_{ij}) / Gra(i) \tag{4}$$

$$D_{ip}_BCA = (\sum_j Byte_{ji} - \sum_j Byte_{ij}) / Gra(i) \tag{5}$$

表 1 检测四元组特征

缩写	元素名称
D_{ip}_PC	地址段内一个目的 IP 地址对应的平均数据包的数量
D_{ip}_BC	地址段内一个目的 IP 地址对应的平均数据包的大小
D_{ip}_PCA	地址段内一个目的 IP 地址的平均双向数据包数量差值
D_{ip}_BCA	地址段内一个目的 IP 地址的平均双向数据包大小差值

3 DDoS 攻击检测流程

SDN 网络中受害目标遭到 DDoS 攻击时，其接收到的数据包目的地址单一，且数据包大小和数量会发生异常变化；与此同时，其进出双向数据包大小和数量差异也会越发明显[12]。根据这一特征，本文提出了 SDN 场景中基于双向流特征的 DDoS 攻击检测方案。

待全网网络地址段划分完成后，为了初始化监控网段范围，系统首先下发交换机初始监控流表项，之后为更加精确检出 DDoS 攻击，系统通过周期性提取网络流四元组特征，运行 DDoS 攻击检测算法，逐次迭代更新监控流表粒度，从而逐步缩小包含受害者的网段大小；一旦确定受害者，系统判定网络正遭受 DDoS 攻击，继而下发命令阻断发往受害者的流量。其整体流程如图 2 所示。

3.1 初始状态网络地址段的划分

通常 DDoS 攻击对于实施平台具有一定的依赖性，因此本文提出以下两点合理假设：a) SDN 网络拓扑是 PoP(Point-Of-Presence)[13]级别的；b)任意两个 PoP 级路由器之间的网络流量经过相同的 SDN 交换机。

初始状态下，检测系统对于网络是否发生 DDoS 攻击是未知的，因此系统需要监控网络中尽可能多的 IP 地址，然而考虑到交换机流表空间的限制和检测实时性要求，系统需要保证所监控的 IP 网段的范围及数量尽可能的小。据 OpenFlow 协议和 SDN 流表规则定义[14]，SDN 网络中每一个 IP 网段都是由相同的网络前缀构成的，因此初始化阶段系统首先需要利用网络前缀匹配规则，将 PoP 入网点内前缀相同的 IP 地址集合划分在同一个 IP 网段内，为防止 IP 集合不能被单独的一条前缀匹配规则匹配，系统需要自动划分 IP 集合直到所有监控 IP 能被完整划分为止。与此同时，为了提取网络流特征，假设存在一源 IP 地址段 R^s 和目的 IP 地址段 R^t，系统需要统计从源 IP 到目的 IP 之间所有的可能路由，即：$\{\langle s,t\rangle, \forall s \in R^s, \forall t \in R^t\}$,鉴于网络流源 IP 与目的 IP 之间的双向性，与之对应的需要统计目的 IP 到源 IP 之间的路由，即：$\{\langle t,s\rangle, \forall t \in R^t, \forall s \in R^s\}$。统计完成后，本文将每一对源、目的 IP 地址对的网络数据包交互，即：$\{R^s, R^t\}$ 称为一道流 f。

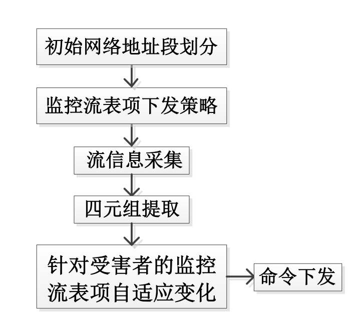


图 2 系统 DDoS 攻击检测整体流程

3.2 交换机监控流表下发策略

3.2.1 监控流表项生成策略

在初始化阶段 IP 地址段划分完成后，为了精确检出 DDoS 攻击，系统需要初始化网络中潜在受害者的 IP 网段范围，同时下发对于网络中每道流 f 的初始监控流表项。一般而言，有两种关于监控流表项的生成策略。第一种策略是对于每一个潜在受害者 IP 网段均设置一条监控流表项。假设现行网络中存在两道网络流：A → B 和 B → A,因此系统会生成 A → B 和 B → A 两条监控规则；但若系统判定 A、B 间皆可能存在受害者进而需要更细粒度生成监控规则时，系统对于地址段 A 会产生 A1 → B，A2 → B,B → A1 和 B → A2 四条监控策略；同时，对于地址段 B 会生成 A → B1,A → B2,B1 → A 和 B2 → A 四条监控规则，故一共生成了 8 条监控规则；而第二种生成策略是仅生成一条监控条目同时负责监控源和目的地址段，一旦源和目的地址段内均被怀疑有受害者，则系统需要更细粒度的划分源和目的地址段。例如当系统想利用第二种生成策略审查地址段 A 和 B 的流量，则会生成 A1 → B1，A2 → B1, A1 → B2, A2 → B2, B1 → A1，B2 → A1，B1 → A2 和 B2 → A2 八条监控条目。

由上述分析可知，若系统为细粒度监控审查地址段 A 和 B 的流量，将其网络范围细分为 k 段，对于第一种监控规则生成策略，针对地址段 A 和 B 总共会产生 4k 条监控条目；而对于第二种生成策略，将总共会产生 $2k^2$ 条监控条目。当 k >2 时，第一种方法将产生更少的监控流表规则。因此，为了尽可能地节省交换机 TCAM 流表空间，本文选用第一种监控流表生成策略。

3.2.2 监控流表项更新策略

系统用算法 1 来迭代更新交换机监控流表项。首先，系统会针对 $\left\{R^s, R^t\right\}$ 中的每道流 f 创建一条单独的监控流表项，即设置 $N_f = 1$。在检测时，系统仅关心可疑受害者 IP 地址段的监控粒度大小。初始化阶段系统将全网监控粒度 G 设置为一个较大的值 G_{upper}，在后续迭代中逐步缩小潜在受害者 IP 范围并逐步减小监控粒度 G；同时，逐次迭代过程系统均需确定网络中所有监控网段最大的监控粒度 $G_{\max}$，并将其审查的地址段范围缩小为原来的一半，随即生成两条监控流表。然而更新流表项粒度后不同 IP 地址段的监控流表项可能会重复，因此为了进一步节省交换机 TCAM 空间需要进行去重操作，逐步更新每道流 f 的监控流表 N_f 。在控制器端计算得到待下发的监控流表项之后，系统会用 3.2.3 节所述的合法性检查算法检验监控流表项下发的合法性。一旦满足要求，控制器便会下发流表项，并进行下一轮迭代过程；否则系统将默认返回上一次迭代过程的监控流表配置，并且最小的监控粒度就是上一轮迭代过程搜索到的 $G_{\max}$。

3.2.3 流表策略下发合法性检查

对于 3.2.2 节监控流表下发而言，最关键的是要考虑所生成的监控流表项所需占用的交换机的空间大小 $\{N_f, \forall f \in F\}$ 是否能得到满足。此类问题可以转换为求解网络流中经典的最大流问题[14]。图 3 形象展示了交换机监控流表项下发合法性检查的数学模型。

模型由四种节点构成，即 *start* 节点 S 、terminate 节点 T 、网络流节点 $\{f_i, i \in F\}$ 和交换机节点 $\{s_j, j \in S\}$ 。 S 节点和所有流节点 f_i 相连并且 S 节点与 f_i 节点之间的链路容量代表 S 到 T 的监控流表项所需占用的交换机 TCAM 流表空间大小。假设正常通信下信道带宽足够， T 到 s_j 代表网络流， s_j 到 T 之间的链路容量代表交换机节点 s_j 剩余的流表空间的大小。由此建模，问题求解便可抽象为： S 到 T 之间的最大流是否等于 $\sum_{f \in F} N_f$，其求解可用 Fold-Fulkerson[15]算法解决，限于篇幅所限，这里不作详述。

算法 1 受害者检测过程监控流表项自适应策略(D)

```
F：网络流集合; R：监控路由规则集合; N_f：对于具体网络流
    f 的监控流表项
Initialize monitor granularity to be G_upper
Initialize monitor rules R for flows in F at granularity G_upper;
    Calculate N_f, for all f ∈ F
if Placement(N_1, N_2, ..., N_F) is not Feasible then
    Raise Error
end if
while 1 do
    G_max = maximum victim range granularity of all rules in R
    R̂ = R
    for all rules r ∈ R̂ with victim range granularity G_max do
        Partition victim IP range into two halves;
        Replace r with two new rules in R̂;
    end for
    Remove redundant rules and Update N_f, for all f ∈ F
     based on R̂;
    if Placement(N_1, N_2, ..., N_F) is not Feasible then
        Return G_max and the previous rule set R
              and the associated allocation
        Break
    end if
    R = R̂;
end while
```

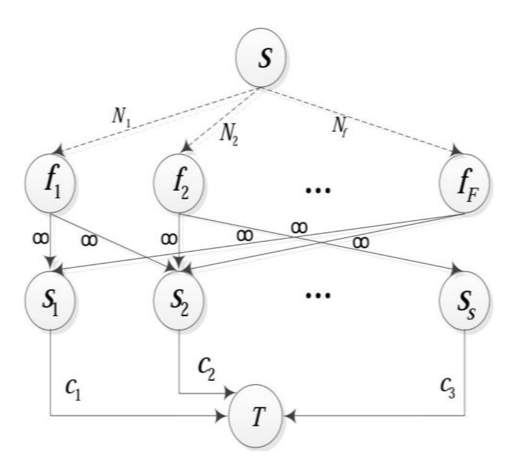


图 3　监控流表项下发最大流问题数学模型

3.3　监控流表项粒度的自适应变化

为了精确检出 DDoS 攻击，本文检测算法需要精确定位受害者位置。然而若受害者网络范围较大，则随着网络通信的推进，交换机流表空间逐渐减小，不利于细粒度监控表项的下发。因此本文引入了一种监控流表项自适应合并或分裂的思想，如图 4 所示。其基本思路为：逐次迭代所产生的监控流表项下发完成后，系统便会周期性提取流表项所审查的 IP 地址段的四元网络特征，并用 GHSOM 算法对其分析分类以检出含有潜在 DDoS 攻击受害者的 IP 地址段。若 GHSOM 判定该地址段内没有 DDoS 攻击受害者，系统便会用粗粒度的监控流表项代替之前细粒度的监控流表项，而对于那些有较高可能性出现受害者的地址段，系统则会建立更加细粒度的监控流表项。

在算法 2 中，本文用迭代的方式逐步缩小潜在受害 IP 的地址段范围。本文利用 Trie 树存储网络监控节点信息，树中每一个节点代表每一个监控流表项中的目的 IP 网段 d 。系统提取

d 的四元流量特征并利用 GHSOM 算法对四元流量特征分析并分类，即计算 $C_v(d)$ 。一旦 $C_v(d)=False$ ，则系统判定 IP 段 d 中没有 DDoS 攻击受害者；而一旦 d 的兄弟节点 $Sib(d)$ 也被判定为 $False$ ，则系统会用其父节点 $Parent(d)$ 的监控流表项代替 d 和 $Sib(d)$ 的监控流表项以节省交换机流表空间。而一旦 $C_v(d)=True$ ，则地址段 d 中可能存在潜在受害者 IP，由此系统会将 d 的监控流表项分裂成 d 的两个子节点的监控流表项以进一步减小监控粒度。同时，由于过大的监控粒度变更将会增大系统检出受害者 IP 的模糊性，进而造成更高的误报率，所以本文规定以 1bit 的大小减小或扩大监控粒度。

算法 2　监控流表项粒度自适应变化算法

```
算法 1 受害者检测过程监控流表项自适应算法(D)
D : R中监控流表项审查的潜在受害者IP地址段
while 1 do
    for all d ∈ D do
        collect packet statistics features ξ(d)
        use victim classifier C_v to calculate C_v(d) to decide
      whether d is under attack or not;
    end for
    for d in D do
        if C_v(d) = False then and C_v(Sib(d)) = False
            Contraction : add monitor rule for victim ranges Parent(d),
      remove monitor rules for victim range d and Sib(d);
        end if
        if C_v(d) = False then and G(d) != 32
            Refinement : add monitor rules for the victim range Child(d),
      remove monitor rules for victim range d;
        end if
    end for
    if Refined Rule Set Infeasible? then
        return list of victim ranges with C_v(d) = True
    end if
end while
```

4　可行性分析与仿真实验

本章主要从检测算法中监控粒度自适应变化的合理性、四元组网络流特征的可行性、DDoS 攻击检测实验结果以及检测算法对于控制器的开销四个方面进行分析。实验中，本文通过 OpenDayLight（Helium）控制器结合 OpenFlow 1.3 开发了 DDoS 攻击检测模块，攻击者与被攻击者网络均由实际主机组成，且用图 4 所示的网络拓扑完成了实验的通信数据交互、采集和 DDoS 攻击检测过程。为了验证本文检测方案在真实网络场景中的 DDoS 攻击检测效果，在部分攻击者主机中安装 LOIC 攻击软件，通过精心构造数据包实际产生了 A 、 N_1 、 N_2 和 N_3 共四种网络流（一种攻击流和三种正常流），具体描述参见表 2。通过随机混合 N_1 、 N_2 和 N_3 来模拟正常网络通信过程中的网络流，且训练与检测分类所用数据参考 WIDE 项目[16],由 ICMP、TCP 和 UDP 三种协议按照 10：80：10 比例混合而成。与此同时，为全面评估检测算法性能，实验中共模拟了三种攻击方式，受害者目标 1、2 和 3 均位于不同网段。方式一攻击的对象仅包含目标 1，方式二攻击的对象包含目标 1 和目标 2，方式三则包含目标 1、目标 2 和目标 3，且三种攻击方式总攻击次数为 2 000 次。

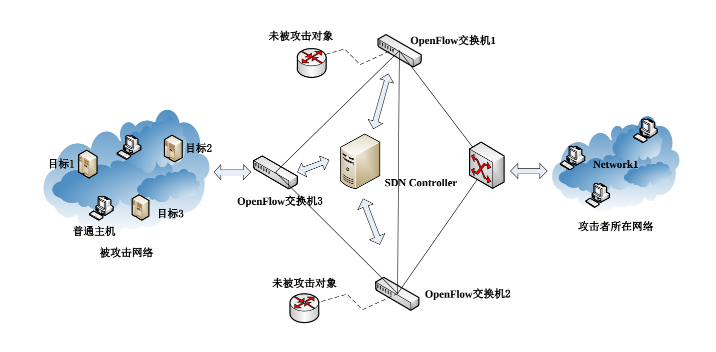


图 4　实验拓扑图

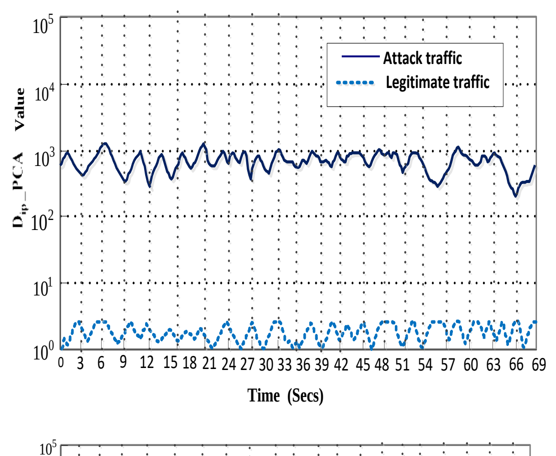
表 2　实验中网络流具体描述

网络流	具体描述
DDoS 攻击流 A	随机选择 N 个源 IP 地址向受害者发送网络包，且保证每个源 IP 地址的发包速率随机化在(30kbps，70kbps）区间中；与此同时，保证收包速率随机化在(1kbps，4kbps）区间中。
高速率正常突发大流量 N_1	从任一源 IP 向另一目的 IP 发送网络包，保证源端发包和收包的速率均随机化在（300mbps，700mbps）之间。
正常非对称小流量 N_2	从任一源 IP 向另一目的 IP 发送网络包，保证源端发包速率随机化在(30kbps，70kbps）区间中，且收包速率随机化在(1kbps，4kbps）区间中，且向同一目的 IP 地址发包的源 IP 数量不超过 50 个。
正常对称小流量 N_3	与 N_2 类似，唯一区别在于源端发包和收包的速率均随机化在(30kbps，70kbps）区间中。

4.1　四元组特征可行性分析及监控粒度自适应改变合理性分析

为分析本文四元组网络流特征的可行性，本节以攻击方式一为例，对检测过程作了详细分析。方式一中，本文选取目标 1 为 DDoS 攻击受害者，从攻击者网络 Network1 中随机选取 100 个 IP 地址作为攻击者。假定包含受害者的 IP 网段为 IP ，其网络前缀长度是 a ，则受害者监控网段可描述为： IP/a 。在实验中，本文对比分析了两种不同的特征组合的检测效果，即：{ D_{ip}_PC , D_{ip}_BC },{ D_{ip}_PC , D_{ip}_BC , D_{ip}_PCA , D_{ip}_BCA }，这两种特征组合区别在于是否考虑发生 DDoS 攻击时的双向流量特征差异。实验中对网络流周期性提取四元组特征，取样周期 Γ 可由系统管理员根据实际情况自行调节，本文设为 3 s。对于 GHSOM 算法的参数设置如下： $\Gamma_m=0.65, T_u=0.03$ ，且学习速率 $\partial(t)$ 设置为 $0.1/(1+0.001t)$ 。正常突发大流量与异常攻击流的区分在以往的检测中是一个难点，因此为了模拟正常突发大流量的场景，实验中适当增大 N_1 在正常流中所占的比例，分别用两种不同

的特征组合作检测分析。本节截取了 DDoS 攻击发生时刻(0-70s)的特征趋势。由于另外 D_{ip}_BC 、D_{ip}_BCA 特征的趋势大致分别与 D_{ip}_PC 与 D_{ip}_PCA 类似，限于篇幅，所以仅展示前两种特征的变化趋势图。从图 5 中清晰可见，突发正常大流量 N_1 与攻击流 A 在单向流特征上区分不明显，而双向流特征的引入使得四特征组合可更好的区分出高速率正常突发大流量 N_1 和攻击流 A，因此四特征组合在受害者检出率上要优于二特征组合，这也一定程度上印证了四特征组合的可行性。



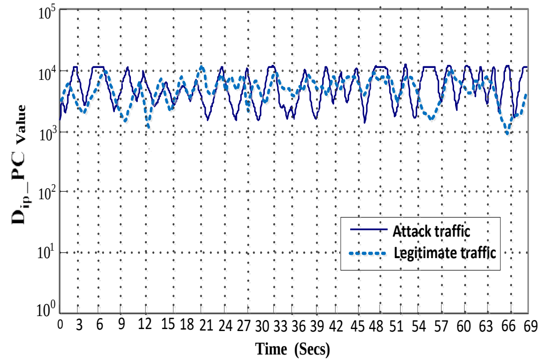


图 5　及 D_{ip}_PCA 特征效果

不同网络前缀下结合上述两种不同特征组合，并利用 GHSOM 算法检测受害者的实验结果如图 6 所示。图中清晰可见，随着监控网络范围的逐步缩小，受害者检出率也在逐步提高。这是因为一旦多道网络流被单一的监控流表项监控，则系统会很难分辨出攻击流 A 和正常非对称小流量 N_2，这也一定程度上印证了本文的监控流表项生成策略的正确性，即：监控粒度越小，检出率越高。同时，从结果上看 GHSOM 在特征分类上效果较好，使用四特征时的 DDoS 检出率明显优于二特征，当使用 4 特征组合且网络前缀长度仅为 16 时，受害者检出率就已大于 80%，并且漏报率为零。

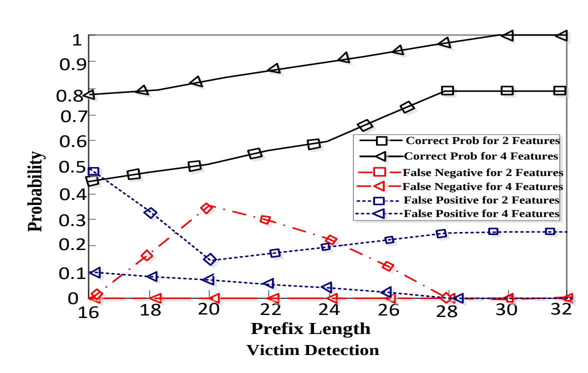


图 6　不同网络前缀下受害者检测效果

4.2　仿真实验结果分析

对于本文 DDoS 检测算法而言，一旦检出受害者，则系统判定此刻网络正遭受 DDoS 攻击，因此受害者的检测准确性十分关键。因此本节主要分析检测算法对于三种攻击方式的检测性能，并与文献[7]的算法过程做对比分析，具体实验结果如表 3 所示。

由表 3 结果可知，本文 DDoS 检测方案在整体检测率上稍优于文献[7]，这主要是由于文献[7]实验拓扑中受害者网络较小，所以在控制器下发监控流表项时总是默认 OpenFlow 交换机 TCAM 表空间充足，监控流表项总是能成功安装于交换机中。而本文为了模拟实际网络环境，受害者网络范围较大(超过 100 台主机且网段分布不均匀)，随着检测算法的迭代运行，交换机 TCAM 流表空间逐渐减小，因此更加精确地定位受害者变得困难。而文献[7]的算法未考虑监控流表项的自适应粒度改变过程，因此在 DDoS 检出率上略差于本文。同时，从实验结果可知，本文的四元组网络流特征可以精确地检出受害者。

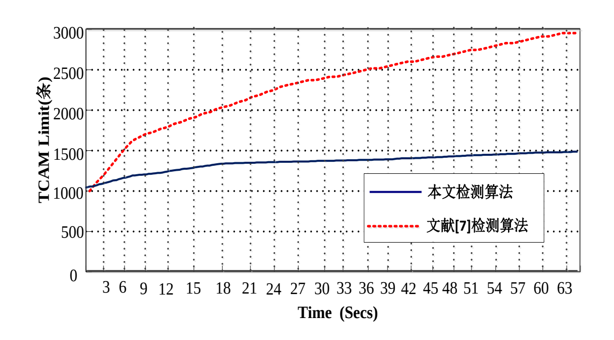
表 3　三种攻击方式受害者检测效果

攻击方式	攻击对象			攻击次数	四元组检测次数	四元组检测率/%	四元组平均检测率/%	文献[7]七元组检测次数	文献[7]七元组检测率/%	文献[7]七元组平均检测率/%
	目标 1	目标 2	目标 3							
方式 1	√			650	627	96.5		606	93.2	
方式 2		√	√	650	623	95.8	96.20	574	88.3	90.7
方式 3	√	√	√	700	675	96.4		634	90.5	

图 7 上图显示了本文检测方法和文献[7]检测方法在交换机 TCAM 表项的变化趋势图。图中清晰可见，8 s 之后，文献[7]的算法对于交换机 TCAM 表项的消耗基本呈线性增长态势，因而随着通信过程的进行，流表项逐渐增多，导致交换机 TCAM 表项空间大幅度减小，不利于后续检测过程的推进；而本文检测方法由于存在监控表表项的粒度自适应过程，因而对交换机 TCAM 表空间消耗相比文献[7]要小很多。

为了尽可能减小误报率，本文的检测方法在逐次迭代过程

中将会消耗一定的交换机流表空间用于更细粒度的监控网络资源。在逐次迭代调整过程中，本文以 1 bit 为单位来调整监控粒度。图 7 下图清晰可见，本文的检测算法对于受害者的检测最终总是可以定位在一台或几台主机之间，检出率较高，且随着算法迭代过程的推进，TCAM 流表空间剩余空间逐渐减小，监控网段的细化也导致后续寻找潜在受害者的迭代次数逐渐减小。综上所述，本文所提出的 DDoS 攻击检测算法在仿真场景中能够以接近 96%的检测率检出 DDoS 攻击，且对于交换机 TCAM 表项的消耗是可接受的。



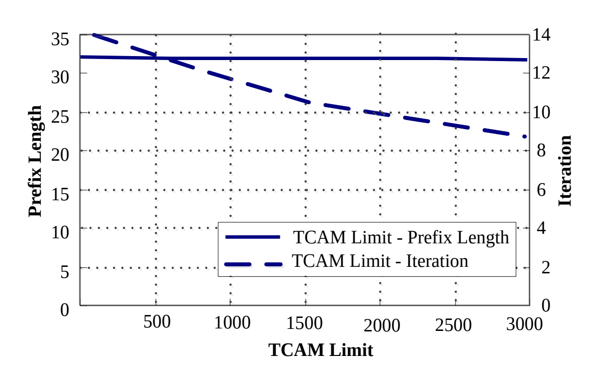


图 7 TCAM 流空间大小和迭代过程关系

4.3 算法开销分析

本文 DDoS 攻击检测算法的开销主要包括训练和检测两个阶段。训练阶段可在线下完成，因此训练过程对于控制器的性能开销可以忽略。检测阶段复杂度主要受待检测目标 IP 地址数量影响，在本实验中以监控网段中具体的主机数量而定。对于每一次检测过程而言，其时间开销受到四元组特征提取和 GHSOM 算法匹配时间影响。四元组特征提取过程涉及前后两次特征数据的比较，若每个周期提取 N 个数据包，则最终处理的数据包共为 $2N$；而 GHSOM 算法的匹配主要受神经元匹配过程影响，其数量远少于受检测目标 IP 数量。对于本文 DDoS 检测算法对于控制器的 CPU 消耗如图 8 所示。

经过多次测试，本文检测算法对于控制器的 CPU 带来的额外开销并不大，在 5%~9%间，为了保证 SDN 网络的安全性，属于可接受的开销范围。

5 结束语

在基于 OpenFlow 的 SDN 网络中，本文研究了如何利用 SDN 控制器集中控制、可编程的优势检出 DDoS 攻击。本文引入了双向流速率特征，更有效区分了大速率正常突发流和 DDoS 攻击流。同时，本文研究了如何通过自适应改变交换机监控流表项粒度以节省 OpenFlow 交换机宝贵的流表空间。仿真实验结果表明，本文的四元组特征及检测方案以接近 96%的 DDoS 高攻击检出率的同时保证了对控制器造成的开销较小。然而由于实验环境的限制，本文实验环境下受害者和攻击者的网络大小仍然无法与实际网络环境相提并论，所以对于未来的研究工作，可集中在以下几点：a)在实际较大规模网络场景中对本文的算法进行验证；b)对于本文而言，监控流表项粒度的自适应调整基本单位为 1 bit，然而为了进一步加快检测速度其粒度大小的自适应动态调整策略值得考虑；c)追踪到受害者之后，考虑利用 SDN 的流表特性进行溯源。

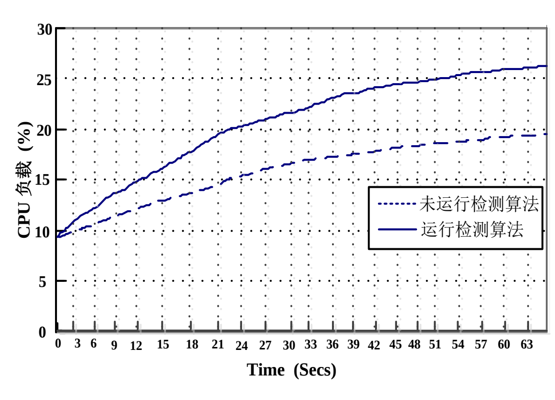


图 8 控制器 CPU 开销图

参考文献：

[1] Ying Z. An adaptive flow counting method for anomaly detection in sdn [C]// Proc of ACM Conference on Emerging Networking Experiments and Technologies. 2013: 25-30.

[2] 2017 年 Q3 全球 DDoS 攻击情况汇总 [J/OL]. (2017-05-17) . http://www. 199it. com/archives/656282. html.

[3] Kandoi R, Antikainen M. Denial-of-service attacks in OpenFlow SDN networks [C]// Proc of IFIP//IEEE International Symposium on Integrated Network Management. 2015: 1322-1326.

[4] Braga R, Mota E, Passito A. Lightweight DDoS flooding attack detection using NOX//OpenFlow [C]// Proc of the 35th Annual IEEE Conference on Local Computer Networks. 2010: 408-415.

[5] Mousavi S M, StHilaire M. Early detection of DDoSattacks against SDN controllers [C]// Proc of International Conference on Computing, Networking and Communications. 2015: 77-81.

[6] Yan Q, Gong Q, Deng F A. Detection of DDoS attacks against wireless SDN controlers based on the fuzzy synthetic evaluation decision-making model [J]. Ad Hoc & Sen-sorWireless Networks, 2016, 33 (1): 275-299.

[7] 姚琳元, 董平, 张宏科. 基于对象特征的软件定义网络分布式拒绝服务攻击检测方法 [J]. 电子与信息学报, 2017, 39 (2): 381-388. (Yao Linyuan, Dong Ping, Zhang Hongke. Distributed denial of service attack detection based on object character in software defined network [J]. Journal of

Electronics & Information Technology, 2017, 39 (2): 381-388.)

[8] 武泽慧，魏强，任开磊，等. 基于 OpenFlow 交换机洗牌的 DDoS 攻击动态防御方法 [J]. 电子与信息学报, 2017, 39 (2): 397-404. (Wu Zehui, Wei Qiang, Ren Kailei, *et al.* Dynamic defense for DDoS attack using OpenFlow-based switch shuffling approach [J]. Journal of Electronics & Information Technology, 2017, 39 (2): 397-404.)

[9] Openflow specification [M/OL]. (2014) http://www. opennet-working. org/images/stories/downloads/sdn-resources/onf-specifications/openflow/openflow-spec-v1. 3. 0. pdf.

[10] Dozono H, Nakakuni M, Kabashima T, *et al.* Analysis of packet traffics and detection of abnormal traffics using Pareto Learning self organizing maps [M]// Neural Infor mation Processing: Models and Applications. Berlin: Springer, 2010: 329-336.

[11] Huang S Y, Huang Y. Network forensic analysis using growing hierarchical SOM [C]// Proc of the 13th IEEE Internation Conference on Data Mining Workshop. 2013: 536-543.

[12] Liu C, Zhang S. A bidirectional-based DDoS detection mechanism [C]// Proc of International Conference on Wireless Communications, Networking and Mobile Computing. 2009: 1-4.

[13] Yan R, Zheng Q, Li H. Combining adaptive filtering and IF flows to detect DDoS attacks within a router [J]. KSII Transactions on Internet & Information Systems, 2010, 4 (3): 428-451.

[14] Sheu R L, Ting M J, Wang I L. Maximum flow problem in the distribution network [J]. Journalof Industrial & Management Optimization, 2017, 2 (3): 237-254.

[15] Ford L R J, Fulkerson D R, Maximal flow through a network [J]. Canadian Journal of Mathematics, 1956, 8 (3): 399–404.

[16] Cho K. MAWI working group traffic archive [EB/OL]. (2016) . http://mawi. wide. ad. jp/mawi/.